

STMIK GI MDP

Program Studi Teknik Informatika Skripsi Sarjana Komputer Semester Genap Tahun 2010/2011

KLASIFIKASI OTOMATIS ARTIKEL PADA WWW.ENSIKLOPEDIA- KOMPUTER.CO.CC DENGAN METODE NEAREST NEIGHBOR

Herman Tansir 2006250019

Abstrak

Tujuan penulisan skripsi ini adalah untuk mengurangi kesalahan pada pemilihan topik. Pada saat input data akan secara otomatis muncul topik. Metodologi penelitian yang digunakan penulis adalah studi literatur dengan cara mencari berbagai informasi tentang Klasifikasi otomatis melalui buku dan web, perencanaan dan analisis berupa pemahaman akan mengenai permasalahan yang dihadapi, perancangan sistem dengan menerapkan metode *Nearest Neighbor* untuk mengklasifikasi artikel pada kelompok *Hardware, Software, Network, Operating System*. Dilakukan pengujian menggunakan database yang terdiri dari 57 artikel dan diperoleh hasil bahwa penggabungan artikel dengan 5 kata kunci lebih efektif dibandingkan 10 kata kunci. Hasil dari aplikasi ini diharapkan dapat membantu pengguna untuk mengklasifikasi topik.

Kata kunci :

Nearest Neighbor , Klasifikasi otomatis, Ensiklopedia Komputer

PERNYATAAN KEASLIAN SKRIPSI

Pernyataan Penyusunan Skripsi

Saya, Herman Tansir

Dengan ini menyatakan bahwa Skripsi yang berjudul :

KLASIFIKASI OTOMATIS ARTIKEL PADA WWW.ENSIKLOPEDIA-KOMPUTER.CO.CC DENGAN METODE NEAREST NEIGHBOR

Adalah benar hasil karya saya dan belum pernah diajukan sebagai karya ilmiah, sebagian atau seluruhnya, atas nama saya atau pihak lain.

Penulis,



Herman Tansir
2006250019

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Fungsi Klasifikasi data yang ada pada *web* dapat mempermudah admin dalam hal klasifikasi data yang diinginkan. Dengan hanya mengisi *Headline*, maka pengguna akan langsung dihadapkan pada data topik.

www.ensiklopedia-komputer.co.cc Adalah sebuah ensiklopedia berbasis *web* yang dibuat dalam Bahasa Indonesia dan hanya membahas seputar artikel dan berita tentang komputer. Ensiklopedia ini terbagi menjadi 5 macam kategori yang dikelompokkan menjadi : *Hardware*, *Software*, *Network*, *Operating System*, dan Serba-Serbi.

Penambahan artikel web www.ensiklopedia-komputer.co.cc saat ini masih dengan menggunakan kode topik secara manual yaitu dengan mengirigira sendiri artikel atau gambar tersebut masuk ke dalam topik *Hardware*, *Software*, *Network*, *Operating System*, dan Serba-Serbi.

Fungsi klasifikasi data akan mempermudah dalam hal penentuan artikel atau gambar yang diinginkan oleh admin. Dengan hanya mengetikkan judul dan memasukkan source dari artikel dan gambar maka artikel dan gambar akan otomatis terklasifikasi ke dalam topik yang bersangkutan.

Melihat keadaan tersebut, maka penulis ingin memperbaiki fungsi penambahan artikel dan gambar pada www.ensiklopedia-komputer.co.cc dengan penentuan topik dilakukan secara otomatis sehingga mempermudah admin yang menggunakan metode *Nearest Neighbor*.

1.2 Permasalahan

Bagaimanakah efektifitas metode *Nearest Neighbor* dalam mengklasifikasikan suatu artikel ke dalam topik-topik yang sesuai pada data di www.ensiklopedia-komputer.co.cc?

1.3 Ruang Lingkup

Pembahasannya di fokuskan pada www.ensiklopedia-komputer.co.cc, untuk beberapa kategori seperti *Hardware, software, network, operating system, serba-serbi* dapat digunakan sebagai search engine.

1.4 Tujuan dan Manfaat

Adapun tujuan serta manfaat dari pembuatan *klasifikasi document* pada Ensiklopedia Komputer ini adalah :

1.4.1 Tujuan Penelitian

Untuk mengukur efektifitas metode *Nearest Neighbor* dalam mengklasifikasikan suatu document ke dalam topik-

topik yang sesuai pada data di www.ensiklopedia-komputer.co.cc.

1.4.2 Manfaat Penelitian

- a. Dengan diperbaikinya fungsi klasifikasi data pada ensiklopedia komputer ini, maka diharapkan dapat memudahkan *admin* dalam mengklasifikasi data.
- b. Mengurangi tingkat kesalahan dalam mengklasifikasi data.

1.5 Metodologi Penelitian

Metodologi Penelitian yang digunakan, yaitu :

1. Studi Litelatur & Pengumpulan Data

Studi litelatur dan pengumpulan data yang digunakan dalam penyusunan laporan dilakukan dengan cara membaca buku-buku, dan jurnal ilmiah yang berhubungan dengan metode pencarian data dan *programming* untuk mendapatkan materi yang dapat dijadikan landasan dan referensi bagi penyusunan skripsi ini. Penulis juga mencari informasi yang berkaitan dengan pencarian data melalui *internet*.

2. Perancangan System

Pada tahap ini, penulis merancang *system* seperti apa, mendeskripsikan system apa yang mau dibuat. Berkenaan dengan system yang ingin dibuat,

maka penulis ingin merancang suatu *system* yang dapat memberikan persamaan kata untuk menentukan topik kepada *user*.

3. Penyusunan Basis Data

Sebelum masuk ke tahap pembuatan program, penulis terlebih dahulu menyusun *database* untuk menampung data di *sever*.

4. Pembuatan Program

Pada tahap ini, penulis membuat program yang akan menangani aktivitas pencarian kata pada data, sehingga mengurangi kesalahan penentuan topik oleh user.

5. Pengujian System

Pada tahap ini, penulis memastikan bahwa situs tersebut benar-benar dapat berjalan dengan baik, atau tidak ada error. Pengujian system dilakukan dengan mencoba mengisi artikel.

6. Perbaikan/Penambahan Data

Apabila terdapat kesalahan setelah pengujian, maka system tersebut akan diperbaiki dan data pada database akan ditambah sehingga ensiklopedia komputer ini semakin lengkap.

7. Analisa dan Simpulan

Pada tahap ini, penulis memuat analisa dan simpulan dari system yang telah selesai diuji.

1.6 Sistematika Penulisan

Sistematika penulisan merupakan gambaran mengenai bab-bab yang disusun oleh penulis dalam laporan skripsi ini. Penulisan skripsi terdiri dari lima bab, dimana tiap bab terdiri dari beberapa sub bab. Susunan garis besar sistematika penulisan skripsi dapat dilihat di bawah ini.

BAB 1 PENDAHULUAN

Bab ini merupakan pendahuluan yang mencakup uraian tentang latar belakang, permasalahan, ruang lingkup, tujuan dan manfaat, metodologi penulisan dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Pada bab ini berisi teori-teori pendukung tentang teori-teori metode *Nearest Neighbor*, yang akan dibahas secara teoritis yaitu dengan memperoleh bahan-bahan dari referensi yang menunjang.

BAB 3 RANCANGAN ALGORITMA DAN PROGRAM

Isi bab ini berisi rancangan penerapan metode *Nearest Neighbor* dalam mengklasifikasikan suatu document ke dalam topik-topik yang sesuai pada data di www.ensiklopedia-komputer.co.cc. beserta perancangan programnya.

BAB 4 IMPLEMENTASI DAN ANALISIS

Pada bab ini dilakukan implementasi dan pengujian terhadap efektifitas pengklasifikasian document ke dalam topic-topik yang sesuai pada data di www.ensiklopedia-komputer.co.cc dengan menggunakan metode *Nearest Neighbor*. Pengujian ini akan dibuktikan kebenarannya, seberapa akurat sistem aplikasi yang dibuat ini.

BAB 5 PENUTUP

Bab ini menjelaskan secara singkat tentang kesimpulan dan saran yang merupakan bagian akhir dari penulisan skripsi ini dan untuk pengembangan sistem selanjutnya.

BAB 5

PENUTUP

5.1 Kesimpulan

Dari hasil penelitian dan pembahasan yang dilakukan, maka dapat disimpulkan bahwa:

1. Metode *Nearest Neighbor* terbukti efektif untuk klasifikasi artikel Ensiklopedia komputer secara otomatis.
2. Pemilihan kata kunci sangat mempengaruhi hasil klasifikasi artikel dengan metode *Nearest Neighbor*, menggunakan perhitungan dengan perbandingan jumlah 5 kata kunci dan 10 kata kunci mempunyai hasil yang sama.
3. Metode *Nearest Neighbor* yang digunakan juga memiliki kelemahan karena perhitungan nilai mempunyai nilai sama maka akan secara otomatis masuk ke dalam topik yang sudah ditentukan.

5.2 Saran

Adapun saran dari penulis yang bisa diterapkan bagi para pembaca yaitu:

1. Fungsi klasifikasi data menggunakan klasifikasi otomatis dengan implementasi *Nearest Neighbor* bisa diterapkan pada fungsi klasifikasi data lain.
2. Klasifikasi otomatis pada fungsi pengklasifikasian topik dapat digabungkan dengan metode lain selain metode *Nearest Neighbor*, agar hasil Klasifikasi bisa lebih mendekati keinginan *user*.

Misal : Support vector Machine, Jaringan Saraf Tiruan,dll.

3. Bagi yang ingin mengembangkan www.ensiklopedia-komputer.co.cc,disarankan agar memperbanyak data artikel dan gambar agar ensiklopedia semakin lengkap.